

**Sequence of a cDNA coding for human glutathione peroxidase confirms TGA encodes active site selenocysteine**

Guy T.Mullenbach\*, Azita Tabrizi, Bruce D.Irvine, Graeme I.Bell<sup>†</sup> and Robert A.Hallewell

Chiron Research Laboratories, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA  
Submitted May 26, 1987

Accession no. Y00369

Selenium dependent glutathione peroxidase (1) is a nuclear encoded cytosolic and mitochondrial enzyme which maintains the integrity of DNA and lipids as well as reducing levels of endogenous hydrogen peroxide as follows:



where ROOH represents peroxidized DNA (2), lipid hydroperoxides, membrane-associated phospholipid hydroperoxides or hydrogen peroxide. We have isolated a cDNA coding the human enzyme from a kidney library in  $\lambda$ gt10 by cross-hybridization with a bovine cDNA (3); 24 of 5300 clones hybridized with the probe. The active site selenocysteine residue (-CH<sub>2</sub>SeH) at position 47 (i.e. SeC) is encoded by the nonsense codon, TGA, as is similarly observed in the mouse gene (4). Interestingly evidence suggests that the selenium atom is incorporated cotranslationally (5) rather than via a posttranslational modification step. This clone possesses 5 bp of the 5'-untranslated region, the 603 bp coding region, 223 bp of the 3'-untranslated region and a canonical polyadenylation signal, AATAAA, upstream of the polyA tract. The amino acid sequence reveals the protein possesses approximately 87% and 85% homology with preprocessed bovine (3) and mouse enzymes, respectively.

Net Cys Ala Ala Arg Leu Ala Ala Ala Ala Glu Ser Val Tyr Ala Phe Ser Ala Arg Pro Leu Ala Gly Gly Glu Pro Val  
 GCCCC ATG TGT CCT GCT CGT CGC CTA CGC CGC CGC CCC GAC TCG GTG TAT GGC TTC TCC CCC CGC CCC CTG CCC CGC GGG GAC CCT GTW 89  
 30  
 Ser Leu Gly Ser Leu Arg Gly Lys Val Leu Leu Ile Glu Asn Val Ala Ser Leu [Ser] Gly Thr Thr Val Arg Asp Tyr Thr Glu Met Asn  
 ACC CTG CGC TCC CTG CGG GGC AAC GTA CTA CTT ATC GAC ATT GTC GGC TCC CTC [TGA] CCC ACC ACC GTC CGG CAC TAC ACC AAC CAG ATG AAC 179  
 40  
 Glu Leu Cln Arg Arg Leu Gly Pro Arg Gly Leu Val Val Leu Gly Phe Pro Cys Asn Cln Phe Gly His Cln Glu Asn Ala Lys Asn Glu  
 GAG CTG CGC CGC CTC CGA CCC CGG GCC CTG GTU CTC CTC CGC TTC CGG TCC AAC CAG TTT CGG CAT CAG AAC CCC AAC AAC CAA 269  
 50  
 60  
 Glu Ile Glu Asn Ser Leu Lys Tyr Val Arg Pro Cty Gly Gly Phe Glu Pro Asn Phe Met Leu Phe Glu Lys Cys Glu Val Asn Gly Ala  
 GAG ATT CAG AAT TCC CTC AAC TAC GTC CGC CCT CGT CGT CGG TTC GAG CCC AAC TTC ATG CTC TTC GAC AAC TCC CAG CTG AAC CCT CGG 359  
 70  
 80  
 Glu Ile Glu Asn Ser Leu Lys Tyr Val Arg Pro Cty Gly Gly Phe Glu Pro Asn Phe Met Leu Phe Glu Lys Cys Glu Val Asn Gly Ala  
 GAG ATT CAG AAT TCC CTC AAC TAC GTC CGC CCT CGT CGT CGG TTC GAG CCC AAC TTC ATG CTC TTC GAC AAC TCC CAG CTG AAC CCT CGG 359  
 90  
 100  
 110  
 Cys Ala His Pro Leu Phe Ala Phe Leu Arg Glu Ala Leu Pro Ala Pro Ser Asp Asp Ile Thr Ala Leu Met Thr Asp Pro Lys Leu Ile  
 CGG CGC CGC CCT CTC TTC CGG CTC CGG GAC CGC CTC CCA CCT CCC ACC AAC CAC GAC GCC ACC CGG CTT ATG ACC AAC CAC CGC AAC CTC ATC 449  
 120  
 130  
 140  
 Cys Ala His Pro Leu Phe Ala Phe Leu Arg Glu Ala Leu Pro Ala Pro Ser Asp Asp Ile Thr Ala Leu Met Thr Asp Pro Lys Leu Ile  
 CGG CGC CGC CCT CTC TTC CGG CTC CGG GAC CGC CTC CCA CCT CCC ACC AAC CAC GAC GCC ACC CGG CTT ATG ACC AAC CAC CGC AAC CTC ATC 449  
 150  
 160  
 170  
 Thr Itp Ser Pro Val \*Cys Arg Asn Asp Val Ala Itp Asn Phe Glu Lys Phe Leu Val Cty Pro Asp Cty Val Pro Leu Arg Arg Tyr Ser  
 ACC TCC TCT CGC CTC TUT CGG AAC GAT GTT GGC TCG AAC ATT GAG AAC TTC CTG GTC CGC CCT GAG CCT GTU CCC CTA CGC AGG TAC ACC AGC 539  
 180  
 190  
 200 201  
 Arg Arg Phe Cln Thr Ile Asp Ile Glu Pro Asp Ile Glu Ala Leu Leu Ser Glu Cty Pro Ser Cty Cys Ala AM  
 CGC CGC TTC CGC ACC ATT GAC ATC GAC GAC CCT GAC ATT GAA CCC CTC CTC TCT CAA CGG CCC AGC TGT CGG TAC CGG CGC CCT TCC TAC CAC 635  
 TCCCAAGTTTCACTGCTCTGCTCGGGCGGTTTCTATGAGGTCTTCTCTAAAGCTTACCGAGGGAGCAACACTCTGATCTTACAGAAAATACCCACCTGGAGATGGGTGCTGGTCC 755  
 TTTCGATCCCAGTCTCTGCCAGAACCCAGGCAAGTTCGGGACTAAATAAACTGGCGGTGTCAGCAAAAAAAAAAAAAA 832  
 Translated Mol. Weight = 21964.60

\*To whom correspondence should be addressed

<sup>+</sup>Present address: Howard Hughes Medical Institute, University of Chicago, Chicago, IL 60637, USA.

## REFERENCES

1. Mannervik, B. (1985) Meth. In Enzym. 113, 490-495.
  2. Christoffersen, B. O. (1969), Biochim. Biophys. Acta 189, 387-389.
  3. Mullenbach, G. T. et al., manuscript in preparation.
  4. Chambers, I., Frampston, J., Goldfarb, P., Affara, N., McBain, W. and Harrison, P. R. (1986) EMBO J. 5, 1221-1227.
  5. Hawke, W. C., Lyons, D. E. and Tappel, A. L. (1982) Biochim. Biophys. Acta 699, 183-191.

Volume 17 Number 5 1989

## Nucleic Acids Research

## Nucleotide sequence of cDNA for rabbit glutathione peroxidase

Masumi Akasaka, Junzo Mizoguchi, Shinichi Yoshimura<sup>1\*</sup> and Keiichi Watanabe<sup>2</sup>

Research Laboratory, Toyo Jozo Co., Ltd., 410-23 Mifuku, Ohito-cho Sizuoka. <sup>1</sup>Cell Biology Research Laboratory and <sup>2</sup>Department of Pathology, Tokai University, School of Medicine, Bohseidai, Ischara-city, Kanagawa, Japan

Submitted January 18, 1989

EMBL accession no. X13837

The cDNA coding rabbit glutathione peroxidase was isolated from liver cDNA library in lambda gt11 by cross hybridization with the rat glutathione peroxidase cDNA which was cloned in this laboratory and reported elsewhere(1). The cDNA consisted of 600 bp of the coding region and the nucleotide sequence revealed that TGA, which is to be the stop codon in general, encoded seleno-cysteine(SeC) residue as was proved to be so with glutathione peroxidases of mouse (2), man(3) and rat(1). The amino acid sequence deduced from cDNA possesses 84, 84, 87 and 85% homology with rat, mouse, human, and bovine(4) enzymes, respectively.

```

1 ATGTGTGCGGCTCGTATGGCGGCGGCTGCCAGTCTGTACTCCTCTCAGCGCACCCGCTGGCCGGCGGGAG
M C A A R M A A A Q S V Y S F S A H P L A G G E
76 CCCGTGAACCTGGGCTCCCTGCAGGGCAAGGTGCTGCTATTGAGAATGTGGCGTCGCTTGAGGCACCTACGGTC
P V N L G S L R G K V L L I E N V A S L [SeC] G T T V
151 CGGGACTACACCCAGATGAACGAGCTGCAAGAGCGCCCTGGGCCCCGGGCCCTGGTCGT$CTGGCTTCCC GTGC
R D Y T O M N E L O E R L G P R A L V V L G F P C
226 AACCAAGTTGGGCATCAGGAGAACGCCAAGAATGAGGAGATTCTGAATTCCCTCAAGTATGTCCGGCCTGGAGGC
N Q F G H Q E N A K N E E I L N S L K Y V R P G G
301 GGTTTCGAGGCCAACTTCATGCTTCCAGAAGTGCAGGGTAACGGCGCAAGGCCAGCCGCTTTGCCTTC
G F E P N F M L F Q K C E V N G A K A S P L F A F
376 CTGGGGAGGGCCCTGCCGCCAGCGACGCCACTGCGCTCATGACCGACCCCAAGTTCATCACCTGGTGC
L R E A L P P S D D P T A L M T D P K F I T W C
451 CCCGTGTGCCGTAAAGCACGTTCTGGAGCTCGAGAAGTTCTGGTGGGGCCCGATGGTGTCCCGTGGCGCAGG
P V C R N D V S W S F E K F L V G P D G V P V R R
526 TACAGCCGCCGCTTCCCCACCATCGACATCGAGGCCGACATCCAAGCCCTGCTGTCCAAGGGGTCTGGCGGTGCC
Y S R R F P T I D I E P D I Q A L L S K G S G G A
601 TAGggcgccccctacccctggctgttgtccggaaaccccccaggtggcgctggctgtgtccatccc 742
* 676 cgaaaaacaaatggagggacgcctgtgtccggaaaccccccaggtggcgctggctgtgtccatccc

```

\*To whom correspondence should be addressed.

## REFERENCES

- Yoshimura, S., Takekoshi, S., Watanabe, K. and Fujii-Kuriyama, Y. (1988) Biochem. Biophys. Res. Commun. 154, 1024-1028.
- Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W. and Harrison, P.R. (1986) EMBO J. 5, 1221-1227.
- Mullenbach, G.T., Tabrizi, A., Irvine, B.V., Bell, C.I. and Hallewell, R.A. (1987) Nuc. Acid. Res. 15, 5484.
- Gunzler, W.A., Steffens, G.J., Grossmann, A., Kim, S.M.A., Otting, F., Wendel, A. and Flohé, L. (1984) Hoppe-Seyler's Z. Physiol. Chem. 365, 195-212.